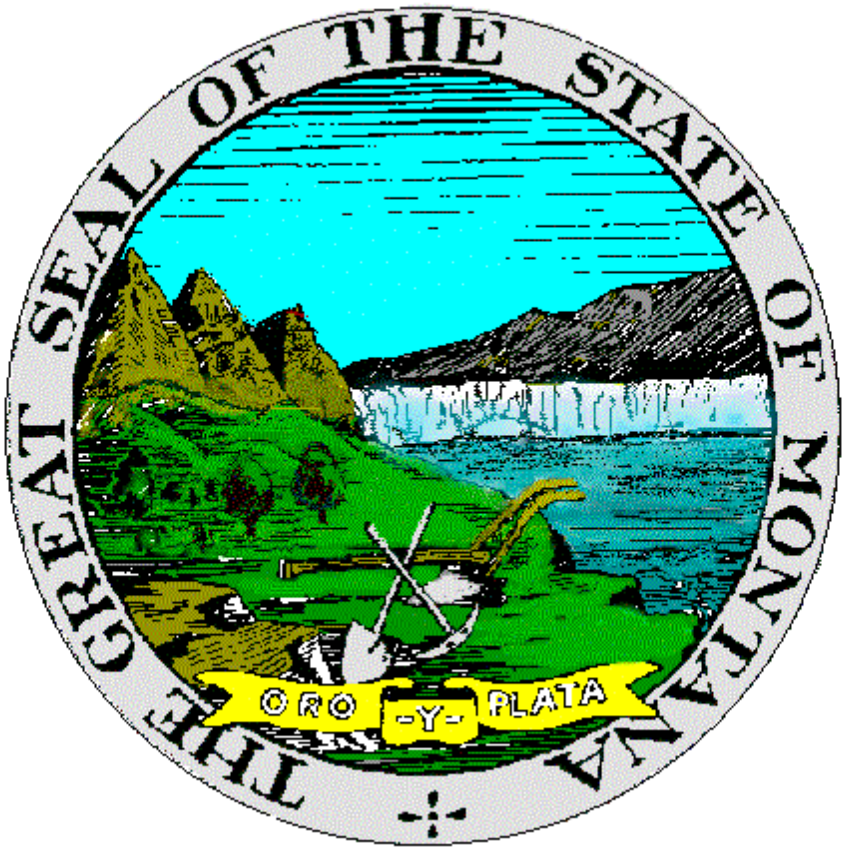
# Office Of Epidemiology And Scientific Support




# DATA DICTIONARY FOR NCHS BRIDGED RACE POPULATION ESTIMATES

# NCHS Bridged Race Population Estimates

The Montana Hospital Discharge Data System (MHDDS) develops datasets based on National Center For Health Statistics (NCHS) Bridged Race Population Estimates. These datasets were developed to be compatible with DPHHS produced numerations of events. Previous estimates were produced upon request from various Montana agencies such as the CEIC or downloaded from the US Census Bureau.

NCHS develops estimates in collaboration with the US Census Bureau, with the addition of race bridging methodology.

Estimates from the US Census Bureau and other agencies were found to be lacking in consistency of form or sufficient granularity for DPHHS needs. For example, the US Census Bureau only releases statewide estimates for population using selected age categories. This is inadequate for estimating teenage pregnancy rates for Montana counties. Rather than rely on pre-selected tabulations, DPHHS uses the fine grained NCHS Bridged Race Population Estimates to produce population estimates as needed.

NCHS Bridged Race Population Estimates provide estimates of mid-year population using US Census Bureau enumeration, vital statistics birth and death records, and IRS and building permits to estimate migration.

# Race Bridging

Race bridging is a technique whereby individuals who select than one race are assigned a primary race for statistical calculation of rates. NCHS originally developed race bridging techniques to provide compatibility between data collected using the 1977 OMB Guidelines (individuals select one primary race from the following supercategories: white, black, Asian / Pacific Islander (PI), and American Indian / Alaska Native) to the 1997 OMB Guidelines (individuals may choose more than one race, division of Asian and PI categories, and expanded literals for capturing race).

In 2003, Montana Office Of Vital Statistics (OVS) implemented the 2003 Revision of the US Standard Death Certificate, which first had racial information collected in accordance with 1997 guidelines. NCHS provides race bridging for these records, which are then merged with OVS demographic information. This process has expanded for Montana birth records (2003 revision implemented in 2008) and fetal death records. For Montana OVS records which are not shared with NCHS (induced abortion records, among others), Montana OVS developed algorithms which parallel NCHS to assign race. Montana OVS also applies these algorithms to records which NCHS codes as being unknown race, based on manual examination that NCHS's algorithms failed to adequately capture information on AI/AN individuals from literal strings.

NCHS applies race bridging methodology to US Census Bureau population data to develop bridged race population estimates. These estimates are otherwise identical to US Census Bureau estimates.

## The NCHS Bridged Race Population Estimates Data Elements

DPHHS's implementation was originally designed for tabulation in the Montana

OVS annual report. Different datasets have been developed based on DPHHS needs. Wherever possible, variable names and definitions are identical between datasets.

Table 1. Variables Included in the NCHS Bridged Race Population Estimates

| | |
|---|---|
| State (MT) and county of residence | Year (1990-Present) |
| Sex | July 1 Population Estimate |
| Age (Single Year) | Hispanic Ancestry |
| Race | Vintage |

Table 2. NCHS Bridged Race Population Estimates Datasets

| Dataset | Description |
|---|---|
| UNIFPOPS | Original dataset |
| DIVERSEPOPS | Expanded version with additional race / Hispanic ancestry categories |
| USUNIFPOPS | US total population dataset |
| USDIVERSEPOPS | US total population dataset with additional race / Hispanic ancestry categories |
| CENUNIFPOPS | Has April 1, 2000 and 2010 Census enumerations |
| CENDIVERSEPOPS | Has April 1, 2000 and 2010 Census enumerations |

DPHHS receives the datasets from a publically available website (http://www.cdc.gov/nchs/nvss/bridged_race.htm). Over the period from 2006-2009, NCHS released estimates as early as June, and as late as October. DPHHS has no influence on the availability of such estimates.

## Build Process

MHDDS downloads the files from NCHS when available. Raw NCHS files are currently downloaded to the directory H:\MHDD\Public\DEMO-POP\NCHS Flat Files. The following processing must be done for the raw files:

0. Files for Montana should be subset to only Montana population estimates.
1. Files must have year transposed.
2. Files must be merged with the 1990s population estimates.
3. Files must have appropriate class variables created.
4. Datasets from NCHS may need to be merged, based on changing NCHS file layouts.

The file NCHS SAS Convert_XXXX.sas in the directory H:\MHDD\Public\DEMO-POP\Maintence Programs does most of the processing.

DPHHS uses the most recent population estimate vintages, which are stored in the final population estimates datasets. Population estimates do change with vintage, the MHDDS recommends that all rates in a table be calculated using the most recent vintage of population estimates, rather than chaining estimates from multiple years together. If estimates for a year are needed before NCHS releases such estimates, the MHDDS recommends using the most recent year available, and informing the reader in a footnote.

Example programs of the use of NCHS population estimates to compute population estimates are in the directory H:\MHDD\Public\DEMO-POP\Example Programs. SAS's PROC TABULATE is ideal for the large number of calculations needed to turn highly granular estimates into usable form.

Table 3. NCHS Bridged Race Population Year Layouts

| Vintage | Description | Dates Of DPHHS Use | Release Date |
|---|---|---|---|
| 2000 Revised | Estimates for 1990-1999 | October 2008-Present (as of Sept 2014) | July 7, 2004 |
| 2005-2007 | Estimates for all years in single flat file | | |
| 2008 | Estimates in 2000-2004 and 2005-2008 in two files. | | |
| 2009 | Separate files for 2000-2009 | Aug 2010-Nov 2011 | |
| 2010 (MTNCHSCEN2010) | Single year estimate based on enumerated population in April 1, 2010 | Nov 2011-Jun 2012 | |
| 2010 | Estimates in 2000-2004 and 2005-2010 in two files | Jun 2012-July 2012 | |
| 2011 | Estimates of 2010-2011 | August 2012-June 2013 | |
| 2010 Revised | Estimates in 2000-2004 and 2005-2010 in two files (Revised Intercensal Estimates) | Nov 2012-Present (as of Sept 2014) | October 26, 2012 |
| 2012 | Estimates of 2010-2012 | June 2013 – June 2014 | June 13, 2013 |
| 2013 | Estimates of 2010-2012 | June 2014 – Present (as of Sept 2014) | June 26, 2014 |

## Directories

NCHS Bridged Race Population Estimates are considered to be public, as they contain no personal identifiers and are provided by NCHS, not Montana DPHHS. As such, they may be copied freely into any directory for any purpose. This is a list of directories which are maintained by members of OESS. Drive letters may vary, depending on mappings.

| Directory | Description |
|---|---|
| | |
| H:\MHDD\Public\DEMO-POP | This mirror is available to users of hospital discharge data. It has raw files from 2000-2009 and for 2012 and 2013. |
| H:\all\NCHS_Population | This mirror should be available to all employees of DPHHS Public Health and Safety division. It does not have the raw data files, as they take much network space. |

## Formats

Many variables have nominal numeric or character values. A format gives the meaning of such values. Formats for variables are given in the directory H:\MHDD\Public\DEMO-POP\Formats, and associated formats for each variable are given in the detail section. For help with formats in SAS, see *My Friend The SAS Format*, by Andrew Karp, available at http://www2.sas.com/proceedings/sugi30/253-30.pdf. For SPSS, value labels are the equivalent of SAS formats.

## Data Use

NCHS Bridged Race Population Estimates are produced by NCHS and publically available. DPHHS makes these estimates available as selected datasets, but is not involved in the original production of such estimates. As such, the datasets produced can be freely posted and distributed. A footnote such as the following is recommended for proper attribution.

| Table 4. Footnotes | |
|---|---|
| | |
| Year | Footnote |
| 2000, 2010 (Census) | 'National Center for Health Statistics.'<br>' Bridged-race estimates of the April 1, 2010 and 2011'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' Available on the Internet at:<br>http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm as of Jan 12, 2003 and Nov 17, 2011' |
| 2010(MTNCHSCEN2010) | 'National Center for Health Statistics.'<br>' Bridged-race intercensal estimates of the July 1, 1990-July 1, 1999; July 1, 2000-July 1, 2009 (Vintage 2009); April 1, 2010'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' prepared by the U.S. Census Bureau with support from the National Cancer Institute.'<br>' Available on the Internet at:<br>http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm as of April 24, 2004; July 23, 2010; November 17, 2011' |
| 2010 | 'National Center for Health Statistics.'<br>' Bridged-race intercensal estimates of the July 1, 1990-July 1, 1999;'<br>' Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2010'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' prepared by the U.S. Census Bureau with support from the National Cancer Institute.'<br>'Available on the Internet at:<br>http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm as of April 24, 2004; May 31, 2012'; |
| 2011 | 'National Center for Health Statistics.'<br>' Bridged-race intercensal estimates of the July 1, 1990-July 1, 1999;'<br>' Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2010; July 1, 2010-July 1 2011'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' prepared by the U.S. Census Bureau with support from the National Cancer Institute.'<br>' Available on the Internet at:<br>http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm as of April 24, 2004; May 31, 2012; July 18, 2012'; |
| 2010 (Revised, 2011) | 'National Center for Health Statistics.'<br>' Bridged-race intercensal estimates of the July 1, 1990-July 1, 1999; July 1, 2000-July 1, 2009'<br>' Postcensal estimates of the resident population of the United States for July 1, 2010-July 1 2011'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' prepared by the U.S. Census Bureau with support from the National Cancer Institute.'<br>' Available on the Internet at:<br>http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm as of April 24, 2004; Oct 26, 2012; July 18, 2012'; |
| 2013 | 'National Center for Health Statistics.'<br>' Bridged-race intercensal estimates of the July 1, 1990-July 1, 1999; July 1, 2000-July 1, 2009'<br>' Postcensal estimates of the resident population of the United States for July 1, 2010-July 1, 2013'<br>' United States resident population by year, county, single-year of age, sex, bridged race, and Hispanic origin,'<br>' prepared by the U.S. Census Bureau with support from the National Cancer Institute.'<br>' Available on the Internet at:<br>http://www.cdc.gov/nchs/nvss/bridged_race.htm as of April 24, 2004; Oct 26, 2012; June 26, 2014'; |

| # | Variable | Type | Length (Bytes) | Label |
|---|---|---|---|---|
| 1 | STFIPS | num | 8 | TWO DIGIT STATE FIPS CODE. DATA ONLY INCLUDES ESTIMATES FOR MONTANA. |
| 2 | CTYFIPS | num | 8 | THREE DIGITS COUNTY FIPS CODE. 1 (BEAVERHEAD) - 111 (YELLOWSTONE) COUNTIES |
| 3 | AGE | num | 8 | SINGLE YEAR OF AGE, EXCEPT FOR 85+ COLLAPSED TO VALUE OF 85 |
| 4 | RACESEX | num | 8 | VARIABLE FOR RACE AND SEX |
| 5 | HISPANIC | num | 8 | HISPANIC ANCESTRY 1 = NO 2 = YES |
| 6 | RACE | char | 5 | RACE - CHARACTER; KEPT FOR LEGACY CODE; 4 VALUES: "WHITE","BLACK","AI/AN","ASIAN" |
| 7 | RACENUM | num | 8 | RACE - NUMERIC; IDENTICAL GROUPINGS AS RACE; 4 VALUES: 1,2,3,4 |
| 8 | SEX | num | 8 | SEX 1 = MALE 2 = FEMALE |
| 9 | YEAR | num | 8 | YEAR = GROUPING VARIABLE |
| 10 | JULPOPCOUNT | num | 8 | POPULATION ESTIMATE FOR JULY 1 OF THE GIVEN YEAR |
| 11 | VINTAGE | Num | 8 | VINTAGE OF POPULATION ESTIMATES |

## FIPS CODE - STATE

State FIPS code allows for identification of the state of population estimates. For the Montana datasets, it is always 30, for national datasets, the variable is not present on the dataset.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| STFIPS | 30, MISSING | None | | |

## FIPS CODE - COUNTY

County of residence is given as three digit numeric codes. In accordance with FIPS guidelines (http://www.itl.nist.gov/fipspubs/fip6-4.htm), Yellowstone Park was eliminated as a separate region in 1990 and incorporated into Gallatin and Park counties.

A table of FIPS codes for Montana counties can be found online at http://mcdc2.missouri.edu/webrepts/commoncodes/ccc_mt.html.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| CTYFIPS | 1-111 | COUNTIES_FIPS.FMT | | |

## AGE

Age is single year of age through age 84, and 85+, which can be tabulated by using an appropriate format.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| AGE | 0-84,85 | AGE_STAND_ADJ.FMT | None | Age is single year of age through age 84, age 85 represents the category of those aged 85 and older. |

## RACESEX

RACESEX is a variable used to derive race and sex. NCHS combines race and sex into a single variable to save space. RACESEX is retained for legacy code.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| RACESEX | 1-8 | | | |

## HISPANIC

HISPANIC is Hispanic ancestry. This variable is available on the DIVERSEPOPS datasets, but not on the UNIFPOPS datasets. Hispanic has unusual coding of 1 for non-Hispanic and 2 for Hispanic.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| HISPANIC | 1,2 | | None | 1 is used for non-Hispanic, 2 is Hispanic. This is backwards from usual coding convention. |

## RACE (Character)

Race is a character variable with values depending on the datasets, providing for four racial categories: white, black, AI/AN, and Asian / PI. It was originally developed for vital statistics data, which was reported using three categories, and eventually expanded to include the original four categories of the datasets. UNIFPOPS is used in many OVS programs, so it is maintained along with the more expansive DIVERSEPOPS. Tabulations not involving race will be identical, regardless of the dataset used.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| RACE (UNIFPOPS) | 'WHITE', 'AI/AN', 'OTHER' | | | |
| RACE (DIVERSEPOPS) | 'WHITE', 'BLACK', 'AI/AN', 'ASIAN' | | | |

## RACE (Numeric)

Race (Numeric) was developed to provide compatibility for merging with datasets that have race coded as a numeric variable, and to provide the ability to use formats to determine.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| RACENUM (UNIFPOPS) | 1,2,9 | | | RACE - NUMERIC; IDENTICAL GROUPINGS AS RACE; 4 VALUES: 1 (White), 2 (Black), 9 (Other) |
| RACENUM (DIVERSEPOPS) | 1,2,3,4 | | | RACE - NUMERIC; IDENTICAL GROUPINGS AS RACE; 4 VALUES: 1 (White), 2 (Black), 3 (AI/AN), 4 (Asian / PI) |

## SEX

SEX is a numeric classification variable.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| SEX | 1,2 | | | 1 for males, 2 for female |

## YEAR

YEAR is the year that the population is estimated for. Note that year is not the same as vintage, the year that a population estimate is created. DPHHS uses updated vintages.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| YEAR | 1990-Present | | | |

## JULPOPCOUNT

JULPOPCOUNT is the variable that gives the estimated midyear population for the race, sex, county, Hispanic ancestry, and year classification. The proper name for this variable should be JULPOPEST, as the variable is an estimate, not a count. However, the use of the count term has stuck in coding, so shall be retained.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| JULPOPCOUNT | Numeric, 0-9999 | | | |

## VINTAGE

VINTAGE is the year of release of estimates. 2004 is used as the vintage year for the final release of the 1990-1999 estimates.

| NCHS BRIDGED RACE POPULATION ESTIMATE VARIABLE NAMES | | | | |
|---|---|---|---|---|
| Current Name | Current Values (2010 forward) | Associated Formats | Legacy Values | Comments |
| | | | | |
| VINTAGE | 2004,2010,2011 | | | |